Adaptive Regularization Strategy for Solving Ill-conditioned Matrices via Zeroing Neural Network

Chenyang Shi
Jishou University, Zhangjiajie, Hunan Province, 427000, China
516230681@qq.com

ABSTRACT

In the fields of scientific computing and engineering control, zero-form neural networks (ZNNs) are widely used for solving time-varying matrices due to their excellent continuous tracking capabilities. However, when the input matrix exhibits highly ill-conditioned characteristics, the basic ZNN and its variants combined with traditional static regularization face an insurmountable contradiction between "solution accuracy" and "numerical robustness." This paper systematically reviews the theoretical bottlenecks of ZNNs in solving ill-conditioned matrices and delves into the evolution and core mechanisms of adaptive regularization strategies. By comparing dynamic strategies based on open-loop time-varying decay mechanisms and closed-loop state error feedback mechanisms, the study points out that the adaptive regulation law driven by real-time residuals can provide strong damping to truncate error divergence in extremely singular transients and automatically fall back when the system returns to a benign state to eliminate steady-state truncation errors caused by excessive smoothing, thus exhibiting excellent global disturbance rejection and tracking performance under complex time-varying conditions. Meanwhile, this paper objectively evaluates the limitations of existing adaptive strategies in terms of hyperparameter empirical tuning, increased single-step computational overhead, and theoretical proof of convergence at predetermined time, and looks forward to the future development trend of constructing rigorous analytical design criteria and exploring lightweight evolutionary architectures.

Keywords: Nullification Neural Network, Ill-Conditioned Matrix, Adaptive Regularization, Error Feedback Mechanism, Numerical Robustness.

1. INTRODUCTION

1.1 Challenges in Solving Ill-conditioned Matrices

In the fields of scientific computing and engineering control, solving time-varying systems often boils down to processing matrix equations. However, in practical applications, we often encounter ill-conditioned matrices with extremely large condition numbers, which are extremely sensitive to small perturbations in the input data. When solving the time-varying equation or generalized inverse, the observation noise or calculation rounding error is amplified, which leads to the numerical solution far from the theoretical solution and even leads to the system deviation [1]. Traditional discrete numerical algorithms perform well in dealing with static good condition matrices, but when they need to solve the inversion problem of time variables and very bad condition matrices, they often have problems because of high computational complexity or lack of continuous tracking ability. Especially in real-time situations, such as dynamic control, traditional algorithms can have digital oscillation [2]. Therefore, it is an important challenge to reduce the error amplification effect while maintaining good real-time performance.

1.2 Current Status of ZNN Applications

Zero-zero neural network (ZNNs) is just like a cyclic neural network for time-varying problems. It uses system error, design dynamics and Lyapunov stability theory to construct a continuous function, and reduces the error exponentially with time. Therefore, they can accurately and continuously track the theoretical solutions that change with time [2]. Compared with the traditional neural network using gradient descent, ZNN converges better when calculating and solving time-varying matrices and nonlinear equations in real time [3]. In recent years, ZNNs have been widely applied in fields such as robotic arm trajectory planning, and enhanced variants with noise suppression capabilities have been derived.

1.3 Significance of Adaptive Regularization Strategy

To address the stability degradation problem of ZNN in solving ill-conditioned matrices, the introduction of regularization techniques has become an important research trend. The traditional Tikhonov method reduces the matrix condition number and suppresses high-frequency noise amplification by adding a regularization term with fixed parameters [1]. However, in the time-varying dynamic environment, the singularity of the bad condition matrix changes with the intensity of external noise. This makes it difficult for static regularization parameters to always keep a proper balance between “solution accuracy and numerical stability”, and it is easy to produce too much smoothing or insufficient regularization [4]. Therefore, it is very important to create an adaptive regularization strategy based on ZNN. This strategy provides a dynamic closed-loop feedback mechanism, which allows regularization parameters to be adjusted independently according to the real-time residual or transient response of the system.

2. ZNN SOLUTION MODEL AND REGULARIZATION PRINCIPLE FOR ILL-CONDITIONED MATRICES

2.1 Basic ZNN Solution Model and its Convergence Characteristics

The theoretical foundation of the ZNN (Zero Neural Network) lies in transforming the complex problem of solving time-varying matrices into the global tracking and dynamic minimization of a continuous residual function by designing specific error evolution dynamics. For the inversion of conventional time-varying non-singular matrices, the basic error function (Zhang function) is defined as the difference between the product of the unknown state matrix and the target matrix and the identity matrix. To ensure that this error matrix decays globally exponentially over time, ZNN derives the core continuous-time evolution dynamic equation based on Lyapunov stability theory:

$$A(t)\dot{X}(t) = -\dot{A}(t)X(t) - \gamma(A(t)X(t) - I) \tag{1}$$

In the above formula, $A(t)$ and $\dot{A}(t)$ represent the known time-varying input target matrix and its corresponding time derivative matrix, respectively; $X(t)$ is the unknown state matrix that the neural network needs to approximate, and $\dot{X}(t)$ is its dynamic update rate; $\gamma > 0$ is the core design parameter that determines the convergence rate of the system error; \$I\$ is the identity matrix of the same order. A large number of theoretical proofs show that, under ideal conditions without external disturbances, the basic model can converge to the theoretical true solution at an exponential speed [5].

2.2 Application of classical regularization techniques in ZNN

To overcome the numerical instability defect of the basic ZNN when dealing with high condition number ill-conditioned matrices, the classical regularization idea in the field of computational mathematics has been introduced into the evolutionary architecture of neural networks. Among them, the Tikhonov regularization method, as a mature singularity correction technique, reconstructs the original ill-conditioned inverse objective into an equivalent “well-conditioned” form. This method no longer directly pursues the absolute zeroing of the original residual equation, but instead optimizes a correction error function with an added smoothing penalty term. The error reconstruction model after introducing Tikhonov regularization can be specifically expressed as:

$$E_R(t) = (A^T(t)A(t) + \beta I)X(t) - A^T(t) \tag{2}$$

In this modified formula, $A^T(t)$ represents the transpose of the original time-varying target matrix; $\beta > 0$ is a pre-defined Tikhonov static regularization parameter (usually a small positive real number); $E_R(t)$ is the new network error state matrix generated after introducing the damping penalty term. Substituting it into the basic ZNN evolution rule, the added damping term βI can forcibly increase the lower bound of the singular values of the ill-conditioned synthesis matrix, thereby strictly constraining its overall condition number within a safe upper bound range [6].

2.3 Performance Bottleneck Analysis of Static Regularization Parameters

Although the introduction of the classic Tikhonov regularization mechanism can effectively alleviate the ill-conditioned sensitivity of the matrix, in actual complex dynamic conditions, directly using the static (i.e., fixed) regularization parameter β exposes a serious performance bottleneck. In actual systems, the ill-conditioned degree of the time-varying matrix and the intensity of external disturbances will fluctuate continuously with the system state. If a large fixed parameter β is set, it can provide enough damping to prevent the system from dispersing to extreme singular state. But in a good state regression matrix, an excessive penalty term will destroy the algebraic structure of the original equation forever. This will cause ZNN to fall into excessive camouflage, and produce stable truncation errors that are difficult to eliminate, which greatly reduces the approximate accuracy of the system [6].

Conversely, if β is set too small in order to balance high accuracy in the well-state stage, the damping provided by this parameter will be severely insufficient when the system encounters short-term extreme ill-conditioned inputs. At that time, due to the amplification effect of conditional inversion, the small rounding error in basic calculation will still multiply, and eventually the neurons in the network will be saturated or collapse directly [7]. Therefore, the fixed and static regularization parameters have a very strong contradiction between the high accuracy of the solution and the strong digital robustness of the neural network. This fixed parameter design itself cannot meet the needs of the time-varying environment, which urges researchers to study the adaptive regularization strategy more deeply.

3. EVOLUTION AND COMPARISON OF ZNN ADAPTIVE REGULARIZATION STRATEGIES

3.1 Dynamic Strategies Based on Time-Varying Decay Mechanisms

To overcome the inherent limitations of static parameters in handling ill-conditioned matrices, early research pioneered the introduction of time-evolution-based dynamic adjustment mechanisms. The core design concept of this strategy is: in the initial stage of system solution, due to the large initial residuals caused by matrix ill-conditioning, the system is prone to numerical divergence. At this time, the neural network must be given a sufficiently large regularization damping to cut off the chain amplification of errors; as time goes by and the solution process progresses, this damping parameter should gradually decrease according to a preset monotonic function law.

Through this open-loop time-varying decay design, the system can effectively suppress numerical oscillations caused by high condition numbers in the initial transient state, and can fundamentally eliminate the over-smoothing phenomenon and steady-state truncation error caused by fixed parameters by gradually approaching zero when approaching steady state [8]. However, such dynamic strategies with preset time trajectories are essentially open-loop control, and their parameter evolution is completely independent of the real-time operating state of the system. Once the parameters decay to a minimum value in the later stages of the solution, if the time-varying target matrix suddenly encounters strong high-frequency disturbances or falls into an extremely singular state again, the regularization term will completely lose its damping protection capability, and small computational errors will still accumulate rapidly and cause the entire neural network solver to collapse [9].

3.2 Adaptive Mechanism Based on State Error Feedback

To overcome the blindness of open-loop time-varying strategies in dealing with sudden disturbances, researchers have introduced closed-loop control theory into the evolutionary architecture of neural networks and proposed an adaptive regularization mechanism based on state error feedback. This mechanism completely abandons the artificially preset time decay function and instead extracts the norm of the dynamic residual matrix (or the transient singular value distribution of the target matrix) in real time during the ZNN evolution process as the driving signal, constructing an independent differential evolution dynamic equation for the regularization parameter.

Under the closed-loop error feedback mechanism, the magnitude of the regularization parameter shows a high positive correlation or dynamic mapping relationship with the real-time residual of the current system [10]. When the pathological condition worsens or external noise intrusion causes a surge in system residuals, the adaptive law can drive the regularization parameter to increase spontaneously in a very short time, thereby providing strong transient damping to cut off error divergence; and when the system gradually converges and returns to a benign environment, the parameter can automatically fall back to a very small safe threshold under the adjustment of the feedback mechanism to ensure that the approximation accuracy is not affected. This two-way fitting network, which uses errors to push parameters and eliminates errors, makes ZNN have strong adaptive disturbance rejection ability in the face of extreme pathological fluctuations [9].

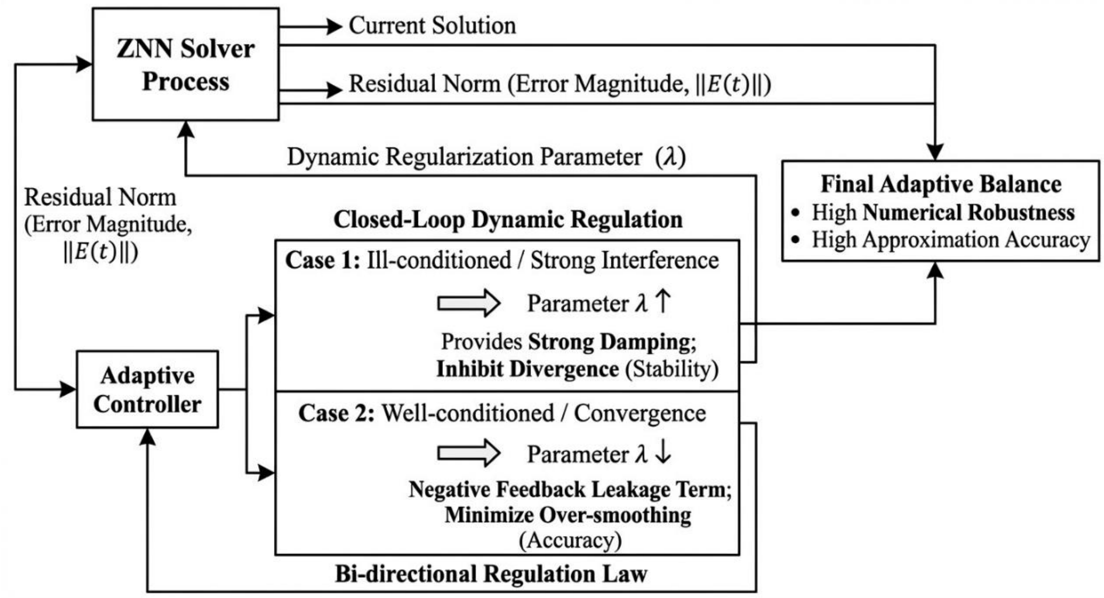

Figure 1. ZNN closed-loop adaptive regularization bidirectional adjustment framework based on state error feedback.

As shown in Figure 1, the adaptive regularization mechanism based on state error feedback makes the regularization parameters respond accurately to the real-time errors of the system by constructing a closed dynamic adjustment cycle. When the solver encounters an unconditional disturbance, which leads to a great improvement in the error standard, the two-way adjustment rule will increase the regularization parameters rapidly, thus providing a powerful way to prevent the numerical deviation of the system. In the well-state stage when the system becomes stable, the parameter itself will decrease due to negative feedback leakage term, so we reduce the excessive fuzzy effect caused by excessive damping.

3.3 Performance Evaluation and Limitations of Various Adaptive Strategies

Combining the results of theoretical convergence analysis and large-scale numerical simulation, the adaptive regularization strategy based on closed-loop error feedback shows that it is much better than static punishment and open-loop time decay method in important things, such as convergence speed, stable approximate accuracy and overall robustness. In complex dynamic experiments, for example, when the inverse dynamics of redundant robot arms need to be solved, the adaptive model with feedback can not only greatly reduce the steady-state residual (several orders of magnitude), but also maintain a smooth and undifferentiated trajectory even if the system has continuous disturbances that change with time [10].

However, current adaptive strategies still face some limitations that urgently need to be addressed when moving towards in-depth engineering applications. First, the tuning process of many hyperparameters involved in the adaptive update law (such as error gain coefficients, negative feedback leakage terms, etc.) still heavily relies on expert rules of thumb and trial and error, and a rigorous analytical design criterion has not yet been established [8]. Second, the introduction of complex closed-loop feedback mechanisms inevitably increases the matrix operation overhead of single-step network evolution, which poses a real-time challenge to the underlying microcontrollers with extremely limited computing resources. Finally, for ZNN systems with variable structure adaptive regularization terms, the theoretical proof of global convergence within a predefined time still faces significant mathematical difficulties, and deeper dynamic analysis tools are urgently needed [9].

4. CONCLUSION

This paper systematically reviews the theoretical bottlenecks of zero-form neural networks (ZNNs) in solving time-varying ill-conditioned matrices and deeply analyzes the evolution and core mechanism of adaptive regularization strategies. Research shows that traditional static regularization suffers from an insurmountable rigid contradiction between “approximation accuracy” and “numerical robustness.” Adaptive strategies, from open-loop decay time-varying to feedback closed-loop state error, can provide dynamic response to extreme singular fluctuations. In particular, the closed-loop dynamic fitting law based on real-time residuals provides strong damping in poor transient conditions to stop error divergence, and automatically reduces errors when the system returns to a benign state to eliminate steady-state truncation errors, thus showing excellent overall tracking performance and disturbance rejection ability under complex dynamic conditions.

Although the adaptive regularization strategy greatly improves the performance of neural network solvers, they still have problems of deep theoretical evidence and technical implementation. Nowadays, it is often necessary to try many things to adjust the important superparameters in the adaptive law, and the complicated nonlinear feedback mechanism makes each step of calculation more arduous, which is very difficult for the speed of the machine. In the future, we must work hard to formulate clear design rules, better prove mathematically how the variable structure system can produce results in a given time, and find a lighter adaptive structure, so that we can use this technology in the fields of redundant dynamic control and real-time signal processing.

REFERENCES

[1] Mourtas, S. D., Katsikis, V. N.: Zeroing neural network for pseudoinversion of an arbitrary time-varying matrix based on singular value decomposition. Mathematics 10(8), 1208(2022).

[2] Petković, M. D., Stanimirović, P. S.: Time-varying pseudoinversion based on full-rank decomposition and zeroing neural networks. Mathematics 11(3), 600(2023).

[3] Xiao, L., Dai, J., Li, Y.: Solving perturbed time-varying linear equation and inequality problem with adaptive enhanced and noise suppressing zeroing. IEEE Transactions on Industrial Informatics 18(11), 7760-7770(2022).

[4] Jin, J., Qiu, J., Zheng, W. X.: Novel adaptive zeroing neural dynamics schemes for temporally-varying linear equation handling applied to arm path following and target motion positioning. Neural Networks 165, 103-115(2023).

[5] Jiang, W., Lin, C. L., Katsikis, V. N., Mourtas, S. D., Stanimirović, P. S., Simos, T. E.: Zeroing neural network approaches based on direct and indirect methods for solving the Yang-Baxter-like matrix equation. Mathematics 10(11), 1950(2022).

[6] Zhang, H., Yin, H.: Zeroing neural network model for solving a generalized linear time-varying matrix equation. AIMS Mathematics 7(2), 2266-2280(2022).

[7] Katsikis, V. N., Mourtas, S. D., Stanimirovic, P. S., Li, X.: A novel zeroing neural network for solving time-varying quadratic matrix equations against linear noises. Mathematics 11(2), 475(2023).

[8] Xiao, L., Zhang, Y., Dai, J.: Comprehensive analysis of a new varying parameter zeroing neural network for time varying matrix inversion. IEEE Access 8, 86604-86614(2020).

[9] Ding, J., Wang, C., Li, C.: A predefined-time adaptive zeroing neural network for solving time-varying linear equations and its application to UR5 robot. IEEE Transactions on Neural Networks and Learning Systems 36(3), 4703-4712(2024).

[10] Xiao, X., Jiang, C., Mei, Q., Zhang, Y.: Noise-tolerate and adaptive coefficient zeroing neural network for solving dynamic matrix square root. CAAI Transactions on Intelligence Technology 9(1), 112-125(2024).